Novel locus influencing retinal venular tortuosity is also associated with risk of coronary artery disease

Abirami Veluchamy[1], Lucia Ballerini[2, 6], Veronique Vitart[3], Katharina E Schraut[4, 5], Mirna Kirin[4,9], Harry Campbell[4], Peter K Joshi[4], Devanjali Relan[6], Sarah Harris[8,10], Ellie Brown[12], Suraj K Vaidya[12], Bal Dhillon[6], Kaixin Zhou[1], Ewan R Pearson[1], Caroline Hayward[3], Ozren Polasek[4,9], Ian J Deary[7,8], Thomas MacGillivray[6,11,12], James F Wilson[3,4], Emanuele Trucco[2], Colin NA Palmer[1*&], Alexander S F Doney[1*&]

[1] Division of Molecular & Clinical Medicine, Ninewells Hospital and Medical School, University of Dundee, Dundee, Scotland, United Kingdom.

[2] VAMPIRE project, Computer Vision and Image Processing Group, School of Science and Engineering (Computing), University of Dundee, Dundee, Scotland, United Kingdom.

[3] MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, Scotland, United Kingdom.

[4] Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, Scotland, United Kingdom.

[5] Centre for Cardiovascular Science, Queen's Medical Research Institute, University of Edinburgh, Edinburgh, Scotland, United Kingdom.

[6] VAMPIRE project, Centre for Clinical Brain Sciences, Royal Infirmary of Edinburgh, Edinburgh, Scotland, United Kingdom.

[7] Psychology, University of Edinburgh, Edinburgh, United Kingdom.

[8] Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, United Kingdom.

[9] Department of Public Health, School of Medicine, University of Split, Split, Croatia.

[10] Medical Genetics Section, Centre for Genomic and Experimental Medicine and MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, Scotland, United Kingdom.

[11] Edinburgh Clinical Research Facility, Western General Hospital, University of Edinburgh, Edinburgh, Scotland, United Kingdom.

[12] Clinical Research Imaging Centre, Queen's Medical Research Institute, University of Edinburgh, Edinburgh, Scotland, United Kingdom.

[&] These authors contributed equally to this work.

* Correspondence authors

c.n.a.palmer@dundee.ac.uk (CNAP), a.doney@dundee.ac.uk (ASFD)

Abstract

Structural variation in retinal blood vessels is associated with global vascular health in humans and may provide a readily accessible indicator of several diseases of vascular origin. Increasing evidence suggests variation in retinal vasculature is highly heritable. This study aimed to identify genetic determinants of retinal vascular traits. We reported a meta-analysis of genome-wide association studies (GWAS) for quantitative retinal vascular traits derived using semi-automatic image analysis of digital retinal photographs from the Genetics of Diabetes Audit and Research in Tayside (GoDARTS) (n=1736) and the Orkney Complex Disease Study (ORCADES) (n=1358) cohorts. We identified a novel genome-wide significant locus at 19q13 (*ACTN4/CAPN12*) for retinal venular tortuosity (*TortV*), and one at 13q34 (*COL4A2*) for retinal arteriolar tortuosity (*TortA*); these two loci were subsequently confirmed in three independent cohorts (n=1413). In the combined analysis in *ACTN4*/*CAPN12* the lead single nucleotide polymorphism (SNP) was rs1808382 (n=4507; Beta=-0.109; standard error (SE) =0.015; $P=2.39\times10^{-13}$) and in *COL4A2* it was rs7991229 (n=4507; Beta=0.103; SE=0.015; $P=4.66\times10^{-12}$). Notably, the *ACTN4*/*CAPN12* locus associated with retinal *TortV* is also associated with coronary artery disease and heart rate. Our findings demonstrate the contribution of genetics in retinal tortuosity traits, and provide new insights into cardiovascular diseases.

Author Summary

Retinal vascular features are associated with wide range of diseases related to vascular health and provide an opportunity to understand early structural changes in vasculature which may help to predict disease risk. Emerging evidence indicates that retinal tortuosity traits are both associated with vascular health and highly heritable. However, the genetic architecture of retinal vascular tortuosity has not been investigated. We therefore performed a genome-wide

association study on retinal arteriolar tortuosity (*TortA*) and retinal venular tortuosity trait (*TortV*) using data from two independent discovery cohorts of 3094 individuals of European-heritage. We found a novel associations at 19q13 (*ACTN4/CAPN12*) for *TortV*, and one at 13q34 (*COL4A2*) for *TortA* at discovery stage and validated in three independent cohorts. A significant association was subsequently found between lead SNPs at 19q13 and coronary artery disease, cardiovascular vascular risk factors and heart rate. We also performed genome-wide association studies for retinal vascular calibres and optic disc radius (*ODradius*) and replicated previously reported locus at 10q21.3 for *ODradius*. Our findings highlight genetic impacts on retinal venular tortuosity and it is association with cardiovascular disease. This may provide a molecular pathophysiological foundation for use of retinal vascular traits as biomarkers for cardiovascular diseases.

Introduction

Retinal vascular traits can be readily measured non-invasively from fundus images and changes in these traits have been linked to a number of clinical conditions associated with vascular health including cardiovascular disease[1,2], stroke[3], hypertension[4], and neurodegenerative disease[5]. The association between retinal vascular calibers and cardiovascular disease has been reported in numerous studies and structural variation in retinal vasculature could predict cardiovascular risk[6–8]. More recently deep learning applied to retinal images has been successfully used to predict cardiovascular risk factors and outcomes[9].

Increasing evidence has shown a significant genetic component to variation in retinal blood vascular traits[10,11]. Understanding the molecular genetic architecture of retinal vascular features provides a molecular pathophysiological basis linking retinal microvascular features with systemic vascular pathology. Recent genome-wide association studies (GWAS)

found number of loci for widely investigated retinal traits including central retinal vein equivalent (*CRVE*)[12–14], retinal arteriolar equivalent (*CRAE*)[12–14], and optic disc morphology[15]. Evidence suggests retinal vascular tortuosity, another potentially important vascular parameter, is also associated with a range of cardiovascular risk factors[16,17]. Heritability estimates for retinal arterial tortuosity range from 50-82% and 21% for retinal venular tortuosity[18,19], indicating a substantial genetic contribution to the variation in these parameters.

To our knowledge, no studies have performed genome-wide scan on retinal vascular tortuosity traits. We therefore carried out a discovery stage GWAS analysis using two independent cohorts including the Genetics of Diabetes Audit and Research in Tayside Study (GoDARTS) and the Orkney Complex Disease Study (ORCADES) to examine the underlying genetic factors influencing the retinal vascular tortuosity traits derived from digital retinal photographs; arteriolar tortuosity (*TortA*), maximum *TortA* (*TortAmax*), venular tortuosity (*TortV*), and maximum *TortV* (*TortVmax*). We also conducted a GWAS of other previously investigated retinal vascular traits including *CRAE, CRVE*, Arteriole-to-Venule ratio (*AVR*), as well as Optic Disc radius (*ODradius*). We confirmed our findings for retinal tortuosity traits in three independent replication cohorts (Lothian Birth Cohort 1936 (LBC1936), CROATIA- Korčula, and CROATIA-Split). Moreover, we examined the relationship between these sentinel SNPs and cardiovascular risk factors using data from the Coronary Artery Disease Genome wide Replication and meta-analysis plus The Coronary Artery Disease (CARDIoGRAMplus C4D) consortium meta-analysis[20], Global Lipid Genetics Consortium analysis[21] (GLGC), International consortium for blood pressure (ICBP) GWAS analysis[22] and on heart rate from the UK Biobank [23].

Results

Study samples

Two independent discovery cohorts were included; patients with type 2 diabetes from the GoDARTS (n=1736) and a population-based sample comprising the ORCADES (n=1358). In both these cohorts, traits were measured from retinal fundus images (**S1 Fig.**) using VAMPIRE 3.1[24,25] (Vascular Assessment and Measurement Platform for Images of Retina), which enables efficient, semi-automatic measurement of the retinal vasculature from large numbers of images. The VAMPIRE methodology used in the discovery stage has been previously reported. [25–27]. The study design and characteristics of the discovery cohorts are shown in **Fig 1.,** and **Table 1**.

Table 1. Descriptive statistics of variables for discovery study cohorts.

Traits	GoDARTS	ORCADES
Sample size	1742	1358
Age (yr.)	69.78 ± 9.63	52.23 ± 14.7
Male/Female	990/752	542/816
***ln*TortA (mean ± SD)**	-9.76 ± 0.99	-10.94 ± 1.32
***ln*TortA max (mean ± SD)**	-8.41± 0.92	-9.22 ± 1.17
***ln*TortV (mean ± SD)**	-10.03± 0.78	-11.52 ± 1.01
***ln*TortV max (mean ± SD)**	-8.66± 0.86	-9.76 ± 0.90
ODradius (mean ± SD)	208.56± 18.97	183.26 ± 19.46
CRAE (mean ± SD)	33.11± 2.39	29.09 ± 2.55
CRVE (mean ± SD)	43.74± 3.11	39.15 ± 3.62
AVR (mean ± SD)	0.76± 0.05	0.75 ± 0.08

Eight retinal vascular parameters are abbreviated as follows: Natural log transformed – Retinal Arteriolar Tortuosity (*ln*TortA), maximum TortA (*ln*TortAmax), Retinal Venular Tortuosity (*ln*TortV), and maximum TortV (lnTortV). Optic Disc radius (Odradius), Central Retinal Arteriolar Equivalent (CRAE), Arteriole-to-Venule ratio (AVR), Central Retinal Venular Equivalent (CRVE), SD – standard deviation. GoDARTS: Genetics of Diabetes Audit and Research in Tayside; ORCADES: Orkney Complex Disease Study.

In the discovery stage, we performed a GWAS using the GoDARTS cohort for each retinal trait separately and tested the additive effect of each variant, adjusted for age, gender and the first three principal components. Similarly, GWAS was performed for the same traits in the ORCADES cohort, using a mixed model to account for kinship and first three principal components as covariates correcting for population structure. As there were statistically significant differences by age, and gender between two discovery cohorts (both age and gender, P<0.0001) the model was adjusted for these in both discovery cohorts. We combined the summary results from these two cohorts for each trait using a fixed effect meta-analysis and the genomic inflation factor is 0.99. **S1 Table** presents the results from the meta-analysis and independent cohort GWAS analysis. Manhattan plots and regions of interest for tortuosity traits are shown in **Fig 2-3.** Manhattan plots for other retinal traits and QQ plots for all traits are shown in**,** and **S2-S4 Figs**.

This analysis revealed one genome-wide significant ($p<5\times10^{-8}$) SNP, rs56399312, associated with *TortA* at 13q34, in *COL4A2* with moderate heterogeneity ($I^2=0.50$); Beta=0.182, SE= 0.032, P= 2.70×10^{-8}, and another SNP rs9515212 near *COL4A2* that was just below the threshold for genome-wide significance; Beta=0.151, SE= 0.028, P= 8.59×10^{-8}. Conditional analysis on the lead SNP indicated that these are not independent signals (**S2 Table**). Two genome-wide significant SNPs were associated with *TortV*, at 19q13 in *ACTN4* (lead SNP rs1808382; Beta=-0.123, SE= 0.022, P= 1.55×10^{-8}; no heterogeneity, $I^2=0.00$), and at 12q24.33 near *TMEM132D* (lead SNP rs73157566; Beta= -0.294, SE= 0.054, P= 4.07×10^{-8}; low heterogeneity, $I^2=0.10$); these associations at both these loci have not been reported previously with any retinal vascular parameters.

Although we replicated previously reported loci for *CRVE*, we did not find any novel genome-wide significant loci for this trait[12],[14]. In addition, we did not replicate any of

148 the previously reported SNPs associated with *CRAE*[13]. Finally, we replicated a previously
149 reported genome-wide significant locus for *ODradius* at 10q21.3 near *PBLD*/*ATOH7* **(**lead
150 SNP rs61854835; Beta= -3.840, SE= 0.575, P= 4.06×10^{-11}**)** and confirmed a number of other
151 loci for this trait [15,28–30] (**S3 Table**).

152 We selected three lead SNPs near *ACTN4*, *TMEM132*D, and *COL4A2*, that reached
153 significance $P \leq 1.07 \times 10^{-07}$ as well as their effect size and direction being similar across the
154 discovery cohorts, as candidates to carry forward for replication, and confirmed these in three
155 independent cohorts comprised of up to 1413 individuals of European ancestry including the
156 Lothian Birth Cohort 1936[31] (LBC1936), Croatia-Korcula, and Croatia-Split. Retinal
157 images from these cohorts had been analyzed by SIVA 3.1[32,33] (Singapore I Vessels
158 Assessment) software to quantify the tortuosity traits. The characteristics of the replication
159 cohorts have been presented in **Table 2**. Two *TortA*-associated SNPs, rs7991229, and
160 rs9515212 in *COL4A2* reached significance (P<0.05) in the LBC1936 and Croatians cohorts
161 but lead SNP (rs56399312) did not replicate in the LBC1936. Two *TortV*-associated SNPs,
162 rs1808382 and rs3786835 in *ACTN4/CAPN12* reached suggestive significance ($P<1 \times 10^{-04}$) in
163 the combined analysis of replication cohorts whereas rs73157566 near *TMEM132D* did not
164 replicate.

165 **Table 2. Descriptive statistics of variables for replication study cohorts.**

Traits	LBC1936	Croatia-Korcula	Croatia-Split
Sample size	644	387	382
Age (yr.)	72.48 ± 0.72	54.28 ± 12.51	49.1 ± 14.23
Male/Female	334/310	122/265	154/228
TortA (mean ± SD)	-9.62 ± 0.235	-9.55 ± 0.26	-9.66 ± 0.25
TortV (mean ± SD)	-9.51 ± 0.242	-9.49± 0.21	-9.62 ± 0.22

166 Natural log transformed - Retinal Arteriolar Tortuosity (*ln*TortA), maximum TortA (*ln*TortAmax), Retinal
167 Venular Tortuosity (*ln*TortV), and maximum TortV (lnTortV). SD – standard deviation. LBC1936: Lothian
168 Birth Cohorts 1936; All Croatia: Croatia island of Korcula+ Croatia Split.

169

170 **Meta-analysis of discovery and replication cohorts**

171 In the overall meta-analysis, only SNPs at 13q34, *COL4A2* (*TortA*) and 19q13,
172 *ACTN4* (*TortV*) were confirmed at genome-wide significance. Although *TortA* associated
173 SNPs rs9515212 and rs7991229 were not genome-wide significant in the discovery meta-
174 analysis, they reached genome-wide significance in the overall meta-analysis and no
175 heterogeneity (I^2=0.00) was observed across different cohorts; $P_{overall}=4.66\times10^{-12}$ and
176 $P_{overall}=4.71\times10^{-12}$, respectively. Whereas the lead SNP in *COL4A2* for *TortA* (rs56399312) at
177 discovery stage did not reach genome-wide significance in overall meta-analysis
178 ($P_{overall}=1.95\times10^{-07}$). For *TortV* the lead SNPs, rs1808382 ($P_{overall}=2.39\times10^{-13}$) and rs3786835
179 ($P_{overall}=3.31\times10^{-13}$) near *ACTN4/CAPN12,* maintained genome-wide significance with no
180 heterogeneity (I^2=0.00). These SNPs are in tight LD and therefore do not represent
181 independent signals. **Table 3** contains the summary statistics from replication cohorts and
182 meta-analysis of these cohorts. Forest plots for lead SNPs in the combined analysis are shown
183 in **Fig 4**.

184 ***TortA*-associated variants**

185 *COL4A2* encodes collagen type IV alpha 2, one of the six subunits of type IV
186 collagens which are major structural components of basement membranes, forming a thin
187 sheet of fibers under the endothelium controlling passage of vasoactive substances. These are
188 conserved across species and C-terminal non-collagenous domains play a role in
189 angiogenesis[34]. Recent GWAS report that common variants around *COL4A2* and *COL4A1*
190 (a paralogue immediately proximal to *COL4A2*, with which it shares a promoter and is co-
191 expressed), are associated with coronary artery calcification[35], arterial stiffness[36], and
192 coronary artery disease[20,37–39] (CAD).

193 **Table 3. Results of discovery, replication and overall meta-analysis for tortuosity traits**

SNP	Chr	BP	Candidate gene	Effect allele (Freq.)	Cohort	BETA	SE	P	Het P (I^2)
TortA									
rs7991229	13	111091995	*COL4A2*	G (0.42)	GODARTS	0.136	0.032	2.30×10^{-05}	
					ORCADES	0.187	0.055	0.000728	
					Stage 1 meta-analysis	**0.098**	**0.018**	**1.07×10^{-07}**	0.75(0)
					LBC1936	0.029	0.014	0.039721	
					All Croatia	0.037	0.009	9.83×10^{-05}	
					Stage 2 meta-analysis	**0.114**	**0.027**	**2.80×10^{-05}**	0.62(0)
rs9515212	13	111087563	*COL4A2*	G (0.42)	**Combined**	**0.103**	**0.015**	**4.66×10^{-12}**	0.65(0)
					GODARTS	0.138	0.032	1.96×10^{-05}	
					ORCADES	0.189	0.055	0.000685	
					Stage 1 meta-analysis	**0.099**	**0.018**	**8.59×10^{-08}**	0.75(0)
					LBC1936	0.029	0.014	0.043685	
					All Croatia	0.038	0.009	8.40×10^{-05}	
					Stage 2 meta-analysis	**0.114**	**0.027**	**2.01×10^{-05}**	0.59(0)
rs56399312	13	111121981	*COL4A2*	C (0.269)	**Combined**	**0.104**	**0.015**	**4.71×10^{-12}**	0.85(0)
					GODARTS	0.186	0.037	6.67×10^{-07}	
					ORCADES	0.168	0.066	0.010473	
					Stage 1 meta-analysis	**0.099**	**0.018**	**2.70×10^{-08}**	0.15(0.5)
					LBC1936	0.025	0.018	0.17722	
					All Croatia	0.025	0.012	0.0386	
					Stage 2 meta-analysis	**0.065**	**0.027**	**0.0149**	0.98(0)
					Combined	**0.088**	**0.015**	**1.95×10^{-07}**	0.07(0.6)
TortV									
rs1808382	19	39151034	*ACTN4*	T (0.475)	GODARTS	-0.125	0.026	2.41×10^{-06}	
					ORCADES	-0.116	0.039	0.0031218	
					Stage 1 meta-analysis	**-0.101**	**0.018**	**1.55×10^{-08}**	0.69(0)
					LBC1936	-0.049	0.014	0.00065	
					All Croatia	-0.033	0.009	0.00064	
					Stage 2 meta-analysis	**-0.129**	**0.027**	**1.35×10^{-06}**	0.87(0)
rs3786835	19	111087563	*ACTN4*	A (0.471)	**Combined**	**-0.109**	**0.015**	**2.39×10^{-13}**	0.61(0)
					GODARTS	-0.126	0.026	1.95×10^{-06}	
					ORCADES	-0.109	0.039	0.004924	
					Stage 1 meta-analysis	**-0.099**	**0.018**	**2.26×10^{-08}**	0.87(0)
					LBC1936	-0.049	0.014	0.000537	
					All Croatia	-0.033	0.009	0.0006	
					Stage 2 meta-analysis	**-0.130**	**0.027**	**1.10×10^{-06}**	0.84(0)
rs73157566	12	129533847	*TMEM132D*	A (0.043)	**Combined**	**-0.109**	**0.015**	**3.31×10^{-13}**	0.53(0)
					GODARTS	-0.289	0.061	3.28×10^{-06}	
					ORCADES	-0.314	0.113	0.005458	
					Stage 1 meta-analysis	**-0.097**	**0.018**	**4.07×10^{-08}**	0.28(0.1)
					LBC1936	-0.061	0.045	0.17165	

	GODARTS	-0.003	0.029	0.9045	
	ORCADES	**-0.027**	**0.027**	0.30795	0.35(0)
	Stage 1 meta-analysis	**-0.075**	**0.015**	**2.61×10^{-06}**	0.08(0.6)
	LBC1936				
	All Croatia				
	Stage 2 meta-analysis				
	Combined				

GODARTS: Genetics of Diabetes Audit and Research in Tayside; ORCADES: Orkney Complex Disease Study; LBC1936: Lothian Birth Cohorts 1936; All Croatia: Croatia island of Korcula+ Croatia Split. Natural log transformed - TortA retinal arteriolar tortuosity, TortV retinal venular tortuosity. Standardized beta estimate (Cohen's d) and SE_beta values are in bold text. Beta: Change in natural log transformed retinal tortuosity traits for each copy of the effect allele; SE: standard error; het P, heterogeneity I^2 index; Het P, P value for heterogeneity. Base position is based on build 37 of the reference genome.

Interestingly, gene expression data from GeneAtlas[40], a human protein-coding transcriptome study validated the high expression of *COL4A2* in retinal micro-vessel endothelial cells (**S5 Fig.**) whereas *COL4A1* is weakly expressed in retina indicating a specific role of *COL4A2* in the retinal vasculature. *TortA*-associated variants near *COL4A2* significantly alter transcription factor binding motifs and have putative effects on transcription as annotated by ENCODE (**S4 Table**). Additionally, expression data from the GTEx database[41] confirmed that these significant SNPs, are associated with the expression of *COL4A2* in heart left ventricle and artery aorta, shown in **S5 Table, S6 Fig.,** and these SNPs are in linkage disequilibrium (LD; r^2=0.99, D'=1).

Lead SNPs associated with *TortA* remained significant after conditioning on the previously reported cardiovascular risk variants in *COL4A2* (rs11617955[20], rs4773144[38], rs9515203[39]) (**S7 Fig.**, **S6 Table**). Conversely the lead SNPs for *TortA* were not associated with coronary artery disease (CAD) and myocardial infarction (MI) risk in the CARDIoGRAMplus C4D consortium meta-analysis[20] (**S7 Table**). Finally, the CAD associated variants specifically in *COL4A1* from the CARDIOGRAMplusC4D were not associated with *TortA*, whereas CAD associated *COL4A2* variants are only weakly associated with *TortA* (**S7-S8 Table**). Retinal vascular tortuosity traits have been previously associated with blood pressure[17,19] which may be therefore link these variants with CAD, however,

we found no evidence for an association between these lead variants and blood pressure in the ICBP GWAS analysis[22] (**S9 Table**).

TortV-associated variants

ACTN4 encodes alpha-actinin 4, a cross-linking protein belonging to the spectrin superfamily and mutations in this gene cause focal segmental glomerulosclerosis in humans. *ACTN2*, a homolog of *ACTN4*, interacts with *ACTN4* and missense mutations in *ACTN2* are linked to a range of cardiac diseases[42]. Annotation by ENCODE[43] indicates that the two genome-wide significant variants (rs1808382, rs3786835) associated with *TortV* near *ACTN4* may have direct regulatory effects as they are located within a DNase I hypersensitivity site and in genomic regions enriched for promoter/enhancer histone marks in heart tissues (**S5 Table**). *ACTN4* and *CAPN12* (calcium-activated neural proteases) overlap by 339 bases at their 3' ends and multi-tissue expression quantitative trait loci (eQTL) analysis confirms that these SNPs in *ACTN4* are associated with mRNA expression of both *ACTN4* and *CAPN12* in aorta, tibial artery, atrial appendage and left ventricle of the heart (**S5 Table, S8 Fig.**). Additionally, this analysis indicates that the T allele at rs1808382 is correlated with lower *ACTN4* (artery aorta; $P=2.1\times10^{-03}$) and this correlation is even stronger with *CAPN12* (artery aorta; $P=2.0\times10^{-07}$). However, while gene expression data using GENEINVESTIGATOR validated the high expression of *ACTN4* in arterial tissue, the highest expression of *CAPN12* appears to be in the hematopoietic system (**S5 Fig.**).

Lead SNPs in *ACTN4* were significantly associated with coronary artery disease in the CARDIoGRAMplus C4D consortium meta-analysis[20] (**S7 Table**) and were associated with CAD risk factors; HDL cholesterol and triglycerides in the GLGC[21], but not associated with blood pressure in the ICBP GWAS analysis[22] (**S9 Table**). Furthermore, we have confirmed the association between *TortV* and top variants in *ACTN4*/*CAPN12* in a sensitivity

241 analysis that only included GoDARTS samples without any cardiovascular events prior to the
242 retinal screening date (**S10 Table**). Moreover, recent meta-analysis of 35 GWAS studies
243 reported the association of SNP (rs11083475) in the *ACTN4* locus with increased resting
244 heart rate[44] which may increase cardiovascular disease risk. This signal is the same as that
245 for *TortV* with strong LD being observed between the lead SNPs for *TortV* and the index SNP
246 for heart rate. Furthermore we found that these SNPs were associated with heart rate in UK
247 Biobank[23] (**S11 Table, S9 Fig.**).

248 ## Discussion

249 In this first GWAS for quantitative retinal vascular tortuosity traits, we found novel
250 loci for retinal arteriolar tortuosity (*COL4A2)* and for retinal venular tortuosity
251 (*ACTN4*/*CAPN12)*, which were replicated in three independent cohorts. Our findings are
252 consistent, non-heterogeneous and have the same direction of effects across all five fairly
253 homogeneous cohorts of European ancestry comprising individuals with and without
254 diabetes, and irrespective of the measurement platform used. Notably, we also identified a
255 genome-wide significant signal at a previously reported locus in/near *ATOH7*/*PBLD* for the
256 optic disc radius and replicated previously identified variants for *CRVE* which validate our
257 retinal traits measurement methods. Together these aspects strongly support the robustness
258 our study design and findings. Power calculations indicated that a sample size of 4507 (stage
259 1 and stage 2), and 3094 (stage 1) and the effect size of 0.10 using Bonferroni correction
260 ($P<5\times10^{-8}$) was adequate to provide 80% statistical power to detect the associations.

261 Previous studies have reported the association between *COL4A2* and CAD but the
262 *TortA*-associated variants in *COL4A2* in the present study are not associated with
263 cardiovascular disease and similarly *COL4A2* variants that are associated with CAD do not
264 appear to be associated with arteriolar tortuosity suggesting that variants in this gene complex

may be involved differentially in the pathophysiology of microvascular and macrovascular diseases. However, more work has to be done to determine the distinct role of genetic variants in *COL4A2*/*COL4A1* in different clinical conditions. In contrast, we found that retinal venular tortuosity-associated variants were associated with coronary artery disease as well as heart rate. Furthermore, the lead variant influences the expression of the *ACTN4*/*CAPN12* genes in the heart tissue. Our sensitivity analyses including samples without CAD prior to the date of acquisition of the measured retinal image indicates that relationship between genetic predictors of retinal venular tortuosity and cardiovascular diseases is not due to reverse causation and demonstrate the robustness of our findings.

Evidence suggests that both retinal arteriolar and venular tortuosity traits are associated with blood pressure and cardiovascular risk factors[17]. However, a recent study from the ORCADES and Croatia-Korcula cohorts reported very weak association between retinal arteriolar tortuosity and blood pressure whereas no evident association between retinal venular tortuosity traits and blood pressure[19]. In this regard, neither the *TortA* nor *TortV*-associated variants were associated with systolic and diastolic blood pressure in the ICBP GWAS analysis thus it seems unlikely that observed associations with CAD or related traits are mediated through blood pressure.

In summary, this first GWAS for retinal arteriolar and venular tortuosity reveals SNPs influencing expression of *COL4A2* and *ACTN4*/*CAPN12* respectively. Our results demonstrate that the *TortA*-associated variants in *COL4A2* are independent of CAD, MI, and blood pressure, and point to a selective role of *COL4A2* rather than *COL4A1* in the retinal vessels. Strikingly, we found *TortV*-associated *ACTN4/CAPN12* SNPs are associated with CAD and heart rate but not associated with blood pressure. However, detailed investigation and functional validation of this new finding is essential to elucidate the causal roles of *ACTN4* and/or *CAPN12* in the observed cardiovascular pathophysiology. These findings

highlight the potential genetic impacts of retinal vasculature to provide new insights into cardiovascular disease.

Materials and Methods

Study participants

Discovery Cohorts

Participants in the discovery phase of this study were obtained from the two independent cohorts, the GoDARTS [45] and the ORCADES. GoDARTS comprises individuals of European-heritage from Tayside, Scotland who provided a sample of blood for genetic analysis and consent to link their genetic information to the anonymized electronic health records. Approval for recruitment to GoDARTS was obtained from the Tayside Committee on Medical Research Ethics. 18,190 individuals were recruited with approximately half having type 2 diabetes at the time of recruitment with the other half being diabetes free. 7,290 individuals currently have genome-wide data for analysis. ORCADES is a family-based study of 2078 individuals aged 16-100 years recruited between 2005 and 2011 in the isolated Scottish archipelago of Orkney[46]. Genetic diversity in this population is decreased compared to Mainland Scotland, consistent with the high levels of endogamy historically. Fasting blood samples were collected and over 300 health-related phenotypes and environmental exposures were measured in each individual. All participants provided written informed consent and the study was approved by Research Ethics Committees in Orkney and Aberdeen.

Replication Cohorts

The LBC1936 comprises 1091 participants who were born in 1936, most of whom took part in the Scottish Mental Survey of 1947. At a mean age of 69.5 years (SD 0.8), between 2004 and 2007, they were recruited to a study to determine influences on cognitive aging[31]. The

CROATIA- Korčula study comprises individuals from the Adriatic island of Korčula, between the ages of 18 and 88. The fieldwork was performed in 2007 in the eastern part of the island, targeting healthy volunteers who underwent complete eye examination and provided their blood sample for genetic analysis from the town of Korčula and the villages of Lumbarda, Žrnovo and Račišće. The Croatia-Split study included inhabitants of the Croatian coastal city of Split, aged 18 to 93. The sampling scheme was similar to Croatia- Korčula, and it took place during 2008 and 2009.

Retinal Vascular Parameters Measurement

Retinal image analysis

Discovery cohorts

Standard digital retinal photographs used for routine diabetic retinopathy screening were obtained from the clinical record in 2,104 participants in GoDARTS. Images of the right eye of usable quality, defined using criteria reported in[25,26], were selected and categorized into two datasets based on the image pixel resolution: GoDARTS dataset 1 (n=788) and GoDARTS dataset 2 (n=1288). Finally, 661 images from the GoDARTS dataset 1, and 1083 images from GoDARTS dataset 2 were included after quality control (QC). 28 individual's images from the GoDARTS were excluded due to inadequate resolution. Standard fundus retinal photographs centred between the macula and optic disc were obtained using digital fundus camera from 1,743 participants in ORCADES. After image processing and QC, 1595 individual's retinal images were used for this study.

VAMPIRE 3.0, was used to measure retinal vascular traits in fundus images from both GoDARTS and ORCADES. The measurement process is organized as a sequence of automatic and manual stages. Manual stages allowed correction of errors made by the automatic software (e.g. vessel labeling as artery or vein) and to minimize their impact on

statistical analysis. Standard protocols were followed to measure the retinal vessel parameters. Briefly, after automatic detection of the optic disc and its radius (*ODradius*), the 6 thickest arterioles and 6 thickest venules appearing in a zone extending out from the optic disc boundary to 2 optic disc diameters were sampled to calculate the median (*TortA*) and maximum (*TortAmax*) arteriolar tortuosity and the median (*TortV*) and maximum (*TortVmax*) venular tortuosity. Central Retinal Artery and Vein Equivalent (*CRAE*, *CRVE)* and the Arteriole-to-Venule ratio (*AVR*) qualify vessel calibers and were measured in a zone 2 to 3 optic disc radii from the center of the optic disc. Among the eight parameters, *TortA, TortAmax, TortV* and *TortVmax* mean values were normalized by natural log transformation for association analysis.

Replication Cohorts

Standard retinal fundus images using digital fundus camera from 1091 individuals from LBC1936 were collected at the recruitment stage and three years later, retinal traits were measured at a subsequent wave of testing using SIVA v3.1[32,33] (Singapore I Vessels Assessment), at a mean age of 72.5 years (SD 0.7). A total of 897 and 976 individual's retinal fundus images centered between the macula and optic disc from Croatia- Korčula and Croatia-Split cohorts were collected using digital fundus camera and retinal traits were quantified using SIVA v3.1[32,33]. SIVA is a semi-automated software which can be used to measure the retinal vascular parameters including retinal vascular tortuosity and vascular caliber from retinal images. After automatic detection of the optic disc, it placed a grid with reference to the center of the optic disc. Then the tortuous vessels were identified and tortuosity traits including *TortA*, and *TortV* were measured using the standard grading protocol by the software; this process was monitored by trained graders and adjusted manually if necessary.

Genotyping, quality control and imputation

Discovery cohorts

GoDARTS samples were genotyped using the Affymetrix 6.0 (n=927) and Illumina Human Omni Express (n=809) platforms. The poor quality variants, samples were excluded based on the quality control (QC) criteria included the following: SNPs call rate < 95%, Hardy–Weinberg equilibrium (HWE) P value < 10^{-6}, sample call rate < 95%, sample relatedness (IBD >0.8), and mismatch between reported and genotypic gender information. QC'd genotype data were imputed using IMPUTE2[47,48] on the basis of 1000 Genome Projects reference panel for all population. Finally, ancestry information of the individuals was derived using EIGENSTRAT[49] and first three principal components (PCs) were used for the association analyses to adjust the population stratification. ORCADES samples were genotyped with either the Illumina HumanHap300 bead chip (n=890) or the Illumina Omni1 (n=304) or Illumina Omni Express bead chips (n=1073). Alleles were called in Bead Studio/Genome Studio (Hap300/Omni) using Illumina cluster files. Subjects were excluded if they fulfilled any of the following criteria: genotypic call rate <98%, mismatch between reported and genotypic sex, unexpectedly low genomic sharing with first or second degree relatives, excess autosomal heterozygosity, or outliers identified by IBS clustering analysis. We excluded SNPs on the basis of minor allele frequency (<0.01/monomorphism), HWE ($P<10^{-6}$), call rate (<97%). Given the very high overlap in SNPs between the two Omni chips, the intersection of QC'd SNPs was used to impute and phase individuals' genotyped on the Omni arrays together, whilst the Hap300 individuals were phased and imputed, separately. Samples were phased using Shapeit v2[50]. Imputation was carried out using IMPUTE2 and the 1,000 genomes reference panel. All ancestries phase1 integrated v3 reference panel, with a secondary reference panel of local exome sequences, sequenced using the Agilent Sure Select All Exon Kit v2.0 and Illumina 100 bp paired end reads (average 30x depth), derived

from 90 ORCADES subjects chosen to optimally represent the haplotypes present. Imputations for the Hap300 and Omni subjects were then combined to form a combined panel of 37.5m SNPs for 2222 subjects[51]. Imputed genotypes for 658, 1078, 1358 individuals from the GoDARTS dataset 1, GoDARTS dataset 2 and ORCADES cohorts, respectively, were used for the three independent GWAS analysis.

Replication Cohorts

LBC1936 samples were genotyped at the Wellcome Trust Clinical Research Facility, Edinburgh, using the Illumina Human 610Quad BeadChip. Individuals were excluded based on unresolved gender discrepancy, relatedness, call rate (≤0.95), and evidence of non-Caucasian descent. SNPs were included if they met the following conditions: call rate ≥ 0.98, minor allele frequency ≥ 0.01, and Hardy-Weinberg equilibrium test with P ≥ 0.001. Imputation to the 1000 Genomes (March 2012 release) reference set was performed using minimac software. A total of 1398 participants from the two independent Croatian replication cohorts were available for the analysis and subjects were genotyped on different genotyping platforms including Illumina CNV370v1 and CNV370-Quadv3 for Croatia-Korčula (n=378), and Illumina CNV370-Quadv3 and IlluminaOmniExpressExome-8v1_A for Croatia-Split (n=376). Samples and markers were excluded based on the following QC metrics; SNPs call rate < 98%, HWE with P value < 10^{-6}, sample call rate < 97%, MAF < 1%, outliers identified by IBS clustering analysis and unresolved gender discrepancy. Imputation was carried out using IMPUTE2 software and 1000G Phase I v3 (March 14, 2012) reference panel.

Statistical analyses

We performed association analyses with each data sets from GoDARTS separately for each of the eight retinal traits using SNPTEST V2.5[47], linear regression assuming an additive genetic model, adjusting for 3 ancestry PCs, age at eye examination and gender.

Subsequently, markers with low imputation quality scores (< 0.4) and minor allele frequency cutoffs (< 0.03) were filtered from each GWAS summary output data separately. Then we performed the meta-analysis using a fixed-effects model in GWAMA[52] with the QC filtered data sets. Association analysis in ORCADES was performed for each of the eight retinal traits, using linear mixed modelling to account for relatedness and assuming an additive genetic model, adjusting for 3 ancestry PCs, age at eye examination and gender, using MMscore in ProbABEL[53]. As in GoDARTS, markers with low imputation quality scores (< 0.4) and minor allele frequency cutoffs (< 0.03) were filtered and meta-analysis was performed with the GoDARTS and ORCADES results using GWAMA. The strand alignment and build check between studies were performed prior to meta-analysis. Also, the genomic inflation factor (λ) was estimated by GWAMA (λ=0.99). All statistical analyses and QCs were performed using SNPTEST v2.5[47], ProbABEL[53], GWAMA[52], PLINK v1.09[54], EIGENSTRAT[49], custom shell scripts, and R scripts. Manhattan plots, Quantile-Quantile plots and forest plots were generated using in-build R scripts, and metafor - R package[55]. Regional plots were generated using the Locus Zoom tool[56] and other data processing was performed using R scripts. Conditional analyses were performed in SNPTEST v2.5 using the genome-wide associated loci in the *COL4A2* region, conditioned on lead SNPs (rs56399312). Also, this new locus was conditioned on previously reported genome-wide significant SNPs (rs4773144, rs11617955, rs9515203) associated with coronary artery disease (CAD).

Sensitivity analyses

We performed an association test with an additive model adjusted for age, gender, and first three principal components in the diabetes cohort (GoDARTS) using 759 samples without any cardiovascular events prior to the retinal screening date.

Replication-analyses

435 The top three SNPs ($P \leq 1.07\times10^{-07}$) near *ACTN4*, *TMEM132*D, and *COL4A2* from the
436 discovery stage for the tortuosity traits were taken forward for examination in three
437 replication cohorts of European ancestry. In the LBC1936 cohort, association analysis was
438 performed for arterial and venular tortuosity traits using linear regression model adjusting for
439 age at eye examination, sex, and 3 ancestry PCs, using mach2qtl. Similarly, in the Croatia –
440 Split, - Korčula cohorts, association analysis were performed for each traits separately using
441 the mixed model in R - hglm package to account for kinship derived using gkin function of
442 the GenABEL package[57].

443 Then we combined the summary association statistics for lead SNPs associated with *TortA*
444 and *TortV* from the two discovery and three replication cohorts and effect estimates from
445 each cohort were presented in the forest plots using metafor - R package. Due to the
446 difference in the units of the beta and standard errors between the discovery and replication
447 studies arising from different approaches to measurement we standardized the effect
448 estimates (using Cohen's d) from each of the individual study cohort.

449 **Power calculation**

450 The statistical power of detecting SNPs associations with the quantitative traits in two stage
451 GWAS was calculated using the GWASPower/QT.

452 ***In-silico* look-ups of the novel variants for clinical outcomes**

453 We performed *in-silico* look-ups of variants of interest for cardiovascular related outcomes
454 including coronary artery disease, myocardial infarction, hypertension, HDL, and
455 triglycerides from the CARDIoGRAMplus C4D consortium[20], GLGC [21] and the
456 ICBPGWAS analysis[22]. The CARDIoGRAMplusC4D 1000 Genomes-based meta-analysis
457 data comprised of 60,801 CAD cases and 123,504 controls from European, South Asian, and
458 East Asian descent. In the GLGC, genetic data from 188,577 individuals of European, East

Asian, South Asian, and African ancestry were used to examine the genetic loci associated with blood lipids levels. The ICBP GWAS investigated the genetic loci associated with systolic and diastolic blood pressure traits in 200,000 individuals of European descent. We retrieved summary association results for the index SNPs from these studies to investigate the association of the lead SNPs for *TortA*, and *TortV* with cardiovascular outcomes.

A recent study reported the association of *ACTN4* locus with heart rate[44]. In order to examine whether the lead SNPs associated with *TortV* in *ACTN4* were also associated with heart rate, we checked the LD (r^2>0.8) between our SNPs and the index SNP (rs11083475) for heart rate in that study. Furthermore, we investigated the association of these SNPs with pulse rate in the UK Biobank data. This data comprised of 112,008 participants who had a measure of pulse rate at the main interview and had genotype data. We extracted the imputed genotypes for these SNPs from the interim release data set of the UK Biobank[23] and performed multiple linear regressions including covariates of age, gender, and the first ten principal components obtained using EIGENSTRAT.

In-silico **functional annotation**

The sentinel genome-wide significant variants were mapped to the gene, 20 kb upstream/downstream using BEDTools[58], and UCSC Genome Browser[59]. Top SNPs were queried in the HaploReg v4.1 database[60] to catalogue the all SNPs near noncoding variants with r^2 > 0.8, and RegulomeDB[61]**,** and GWAS catalog databases[62] used to explore the known and predicted regulatory elements and relevant genetic association studies. Functional effects of the top genes were predicted using the Encyclopedia of DNA Elements[43] (ENCODE) project and Roadmap Epigenomics projects which aggregate the information about the transcription factor, motifs, histone modification, and chromatin states. Additionally, functional elements were investigated using HaploReg, UCSC Genome

Browser, and RegulomeDB. We used the expression Quantitative Trait Loci (eQTL) browser database in Genotype-Tissue Expression[41] (GTEx) to examine the cis-eQTLs for the top retinal traits associated SNPs mapped to the gene within the genomic region. Gene Visible web database from GENEINVESTIGATOR which integrates manually curated gene expression data from microarray and RNAseq experiments, was used to find the expression level of the genes, associated with tortuosity traits, in the human tissues.

Acknowledgements

We are grateful to all the participants in the GoDARTS study, the general practitioners, the Scottish School of Primary Care for their help in recruiting the participants, and to the whole team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. The study complies with the Declaration of Helsinki. We acknowledge the support of the Health Informatics Centre, University of Dundee for managing and supplying the anonymized data and NHS Tayside, the original data owner. The Wellcome Trust United Kingdom Type 2 Diabetes Case Control Collection (GoDARTS) was funded by The Wellcome Trust (072960/Z/03/Z, 084726/Z/08/Z, 084727/Z/08/Z, 085475/Z/08/Z, 085475/B/08/Z) and as part of the EU IMI-SUMMIT program.

ORCADES was supported by the Chief Scientist Office of the Scottish Government (CZB/4/276, CZB/4/710), the Royal Society, the MRC Human Genetics Unit, Arthritis Research UK and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We would like to acknowledge the invaluable contributions of the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney.

We thank the Lothian Birth Cohort 1936 (LBC1936) participants and team members who contributed to these studies. Phenotype collection was supported by Age UK (The Disconnected Mind project). Genotyping was funded by the BBSRC (BB/F019394/1). The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC and Medical Research Council (MRC) is gratefully acknowledged.

The Croatia- Korčula and Croatia-Split study were funded by grants from the Medical Research Council (UK), European Commission Framework 6 project EUROSPAN (Contract No. LSHG-CT-2006-018947), European Commission Framework 7 project BBMRI-LPC (FP7 313010), the Republic of Croatia Ministry of Science, Education and Sports research grant (216-1080315-0302) and the Croatian Science Foundation (grant 8875). We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to the University of Split and Zagreb Medical Schools and Croatian Institute for Public Health. The SNP genotyping for the Korčula cohort was performed in Helmholtz Zentrum München, Neuherberg, Germany.

VAMPIRE team: Parts of the VAMPIRE software and its use for measuring the image set described here was funded by the Leverhulme Trust project RPG-419 “Discovery of retinal biomarkers for genetics with large cross-linked data sets”. VAMPIRE 3.1 has been developed under funding from EPSRC (EPSRC EP/M005976/1), the EU (REVAMMAD ITN).

For the analysis of the association of the identified genetic variants with heart rate, this research has been conducted using the UK Biobank Resource under Application Number 20405.

References

1. Flammer J, Konieczka K, Bruno RM, Virdis A, Flammer AJ, Taddei S. The eye and the heart. Eur Heart J. 2013;34(17):1270–8.

2. Liew G, Mitchell P, Rochtchina E, Wong TY, Hsu W, Lee ML, et al. Fractal analysis of retinal microvasculature and coronary heart disease mortality. Eur Heart J. 2011;32(4):422–9.

3. Baker ML, Hand PJ, Wang JJ, Wong TY. Retinal signs and stroke: Revisiting the link between the eye and brain. Stroke. 2008;39(4):1371–9.

4. Ding J, Wai KL, McGeechan K, Ikram MK, Kawasaki R, Xie J, et al. Retinal vascular caliber and the development of hypertension: a meta-analysis of individual participant data. J Hypertens. 2014;32(2):207–15.

5. MacCormick IJ, Czanner G, Faragher B. Developing retinal biomarkers of neurological disease: an analytical perspective. Biomark Med. 2015;9(7):691–701.

6. Nägele MP, Barthelmes J, Ludovici V, Cantatore S, von Eckardstein A, Enseleit F, et al. Retinal microvascular dysfunction in heart failure. Eur Heart J. 2017;(February):47–56.

7. Wang JJ, Liew G, Klein R, Rochtchina E, Knudtson MD, Klein BEK, et al. Retinal vessel diameter and cardiovascular mortality: Pooled data analysis from two older populations. Eur Heart J. 2007;28(16):1984–92.

8. Wang SB, Mitchell P, Liew G, Wong TY, Phan K, Thiagalingam A, et al. A spectrum of retinal vasculature measures and coronary artery disease. Atherosclerosis [Internet]. Elsevier Ltd; 2018;268:215–24. Available from: https://doi.org/10.1016/j.atherosclerosis.2017.10.008

554 9. Poplin R, Varadarajan A V, Blumer K, Liu Y, McConnell M V, Corrado GS, et al.
555 Prediction of cardiovascular risk factors from retinal fundus photographs via deep
556 learning. Nat Biomed Eng. 2018;2(3):158–64.

557 10. Sun C, Zhu G, Wong TY, Hewitt AW, Ruddle JB, Hodgson L, et al. Quantitative
558 genetic analysis of the retinal vascular caliber: The australian twins eye study.
559 Hypertension. 2009;54(4):788–95.

560 11. Liu YP, Kuznetsova T, Jin Y, Thijs L, Asayama K, Gu YM, et al. Heritability of the
561 retinal microcirculation in flemish families. Am J Hypertens. 2013;26(3):392–9.

562 12. Kamran Ikram M, Xueling S, Jensen RA, Cotch MF, Hewitt AW, Ikram MA, et al.
563 Four novel loci (19q13, 6q24, 12q24, and 5q14) influence the microcirculation In vivo.
564 PLoS Genet. 2010;6(10):1–12.

565 13. Sim X, Jensen RA, Ikram MK, Cotch MF, Li X, MacGregor S, et al. Genetic Loci for
566 Retinal Arteriolar Microcirculation. PLoS One. 2013;8(6):1–12.

567 14. Jensen RA, Sim X, Smith AV, Cheng C, Brody JA, Cotch MF, et al. Novel Genetic
568 Loci Associated with Retinal Microvascular Diameter. Circ Cardiovasc Genet.
569 2016;9(1):45–54.

570 15. Macgregor S, Hewitt AW, Hysi PG, Ruddle JB, Medland SE, Henders AK, et al.
571 Genome-wide association identifies ATOH7 as a major gene determining human optic
572 disc size. Hum Mol Genet. 2010;19(13):2716–24.

573 16. Owen CG, Rudnicka AR, Nightingale CM, Mullen R, Barman SA, Sattar N, et al.
574 Retinal arteriolar tortuosity and cardiovascular risk factors in a multi-ethnic population
575 study of 10-year-old children; The child heart and health study in England (CHASE).
576 Arterioscler Thromb Vasc Biol. 2011;31(8):1933–8.

17. Cheung CY-L, Zheng Y, Hsu W, Lee ML, Lau QP, Mitchell P, et al. Retinal vascular tortuosity, blood pressure, and cardiovascular risk factors. Ophthalmology. Elsevier Inc.; 2011;118(5):812–8.

18. Taarnhøj NCBB, Munch IC, Sander B, Kessel L, Hougaard JL, Kyvik K, et al. Straight versus tortuous retinal arteries in relation to blood pressure and genetics. Br J Ophthalmol. 2008;92:1055–60.

19. Kirin M, Nagy R, MacGillivray TJ, Polašek O, Hayward C, Rudan I, et al. Determinants of retinal microvascular features and their relationships in two European populations. J Hypertens. Lippincott Williams & Wilkins; 2017 Aug 13;35(8):1646–59.

20. Nikpay M, Goel A, Won H-H, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet. 2015 Oct 7;47(10):1121–30.

21. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet. 2013;45(11):1274–83.

22. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, et al. Genetic Variants in Novel Pathways Influence Blood Pressure and Cardiovascular Disease Risk. Nature. 2011 Oct 6;478(7367):103–9.

23. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLoS Med. 2015;12(3):1–10.

24. Trucco E, Giachetti A, Ballerini L, Relan D, Cavinato A, MacGillivray T.

Morphometric measurements of the retinal vasculature in fundus images with VAMPIRE. Biomed Image Underst Methods Appl. 2015;91–111.

25. MacGillivray TJ, Cameron JR, Zhang Q, El-Medany A, Mulholland C, Sheng Z, et al. Suitability of UK Biobank retinal images for automatic analysis of morphometric properties of the vasculature. PLoS One. 2015;10(5):1–10.

26. Giachetti A, Trucco E. Accurate and reliable segmentation of the optic disc in digital fundus images the optic disc in digital fundus images. J Med Imaging. 2014;1(2).

27. Lupaşcu CA, Tegolo D, Trucco E. Accurate estimation of retinal vessel width using bagged decision trees and an extended multiresolution Hermite model. Med Image Anal. 2013;17(8):1164–80.

28. Ramdas WD, van Koolwijk LME, Ikram MK, Jansonius NM, de Jong PTVM, Bergen AAB, et al. A genome-wide association study of optic disc parameters. PLoS Genet. 2010;6(6):1–12.

29. Khor CC, Ramdas WD, Vithana EN, Cornes BK, Sim X, Tay WT, et al. Genome-wide association studies in Asians confirm the involvement of ATOH7 and TGFBR3, and further identify CARD10 as a novel locus influencing optic discarea. Hum Mol Genet. 2011;20(9):1864–72.

30. Springelkamp H, Mishra A, Hysi PG, Gharahkhani P, Höhn R, Khor C-C, et al. Meta-analysis of Genome-Wide Association Studies Identifies Novel Loci Associated With Optic Disc Morphology. Genet Epidemiol. 2015 Mar 28;39(3):207–16.

31. Deary IJ, Gow AJ, Pattie A, Starr JM. Cohort profile: The lothian birth cohorts of 1921 and 1936. Int J Epidemiol. 2012;41(6):1576–84.

32. Sasongko MB, Wang JJ, Donaghue KC, Cheung N, Benitez-Aguirre P, Jenkins A, et

623 al. Alterations in retinal microvascular geometry in young type 1 diabetes. Diabetes
624 Care. 2010;33(6):1331–6.

625 33. Koh V, Cheung CYL, Zheng Y, Wong TY, Wong W, Aung T. Relationship of retinal
626 vascular tortuosity with the neuroretinal rim: The singapore malay eye study. Investig
627 Ophthalmol Vis Sci. 2010;51(7):3736–41.

628 34. Kuo DS, Labelle-Dumais C, Gould DB. Col4a1 and col4a2 mutations and disease:
629 Insights into pathogenic mechanisms and potential therapeutic targets. Hum Mol
630 Genet. 2012;21(R1).

631 35. Odonnell CJ, Kavousi M, Smith A V., Kardia SLR, Feitosa MF, Hwang SJ, et al.
632 Genome-wide association study for coronary artery calcification with follow-up in
633 myocardial infarction. Circulation. 2011;124(25):2855–64.

634 36. Tarasov K V. COL4A1 Is Associated With Arterial Stiffness By Genome Wide
635 Association Scan. 2009;2(2):151–8.

636 37. Schunkert H, Konig IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, et al. Large-
637 scale association analysis identifies 13 new susceptibility loci for coronary artery
638 disease. Nat Genet. 2011;43(4):333–8.

639 38. Yang W, Ng FL, Chan K, Pu X, Poston RN, Ren M, et al. Coronary-Heart-Disease-
640 Associated Genetic Variant at the COL4A1/COL4A2 Locus Affects
641 COL4A1/COL4A2 Expression, Vascular Cell Survival, Atherosclerotic Plaque
642 Stability and Risk of Myocardial Infarction. PLoS Genet. 2016;12(7):1–15.

643 39. Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, et al.
644 Large-scale association analysis identifies new risk loci for coronary artery disease.
645 Nat Genet. 2013;45(1):25–33.

646 40. Su AI, Wiltshire T, Batalov S, Lapp H, Ching K a, Block D, et al. A gene atlas of the
647 mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci U S A.
648 2004;101(16):6062–7.

649 41. Ardlie KG, Deluca DS, Segre A V., Sullivan TJ, Young TR, Gelfand ET, et al. The
650 Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in
651 humans. Science (80-). 2015;348(6235):648–60.

652 42. Chiu C, Bagnall RD, Ingles J, Yeates L, Kennerson M, Donald J a., et al. Mutations in
653 Alpha-Actinin-2 Cause Hypertrophic Cardiomyopathy. A Genome-Wide Analysis. J
654 Am Coll Cardiol. Elsevier Inc.; 2010;55(11):1127–35.

655 43. Myers RM, Stamatoyannopoulos J, Snyder M, Dunham I, Hardison RC, Bernstein BE,
656 et al. A user's guide to the Encyclopedia of DNA elements (ENCODE). PLoS Biol.
657 2011;9(4).

658 44. den Hoed M, Eijgelsheim M, Esko T, Brundel BJJM, Peal DS, Evans DM, et al.
659 Identification of heart rate-associated loci and their effects on cardiac conduction and
660 rhythm disorders. Nat Genet. 2013;45(6):621–31.

661 45. Hébert HL, Shepherd B, Milburn K, Veluchamy A, Meng W, Carr F, et al. Cohort
662 Profile: Genetics of Diabetes Audit and Research in Tayside Scotland (GoDARTS). Int
663 J Epidemiol. 2017;(December):1–12.

664 46. McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc
665 L, et al. Runs of Homozygosity in European Populations. Am J Hum Genet.
666 2008;83(3):359–72.

667 47. Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nat
668 Rev Genet. Nature Publishing Group; 2010;11(7):499–511.

669 48. Howie B, Marchini J, Stephens M. Genotype Imputation with Thousands of Genomes.
670 G3. 2011;1(6):457–70.

671 49. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick N a, Reich D. Principal
672 components analysis corrects for stratification in genome-wide association studies. Nat
673 Genet. 2006;38(8):904–9.

674 50. O'Connell J, Gurdasani D, Delaneau O, Pirastu N, Ulivi S, Cocca M, et al. A General
675 Approach for Haplotype Phasing across the Full Spectrum of Relatedness. PLoS
676 Genet. 2014;10(4).

677 51. Joshi PK, Prendergast J, Fraser RM, Huffman JE, Vitart V, Hayward C, et al. Local
678 Exome Sequences Facilitate Imputation of Less Common Variants and Increase Power
679 of Genome Wide Association Studies. PLoS One. 2013;8(7).

680 52. Magi R, Morris AP. GWAMA: software for genome-wide association meta-analysis.
681 BMC Bioinformatics. 2010;11(ii):288.

682 53. Karssen L, Kooyman M, Aulchenko Y, Struchalin M. ProbABEL. 2016;1–24.

683 54. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ, et al. Second-
684 generation PLINK: rising to the challenge of larger and richer datasets. Gigascience.
685 2015;4(1):7.

686 55. Team RC. R: A language and environment for statistical computing. R Found Stat
687 Comput Vienna, Austria. 2014;

688 56. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al.
689 LocusZoom: Regional visualization of genome-wide association scan results.
690 Bioinformatics. 2011;27(13):2336–7.

691 57. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: An R library for

genome-wide association analysis. Bioinformatics. 2007;23(10):1294–6.

58. Quinlan AR. BEDTools: The Swiss-Army tool for genome feature analysis. Current Protocols in Bioinformatics. 2014. 11.12.1-11.12.34 p.

59. Speir ML, Zweig AS, Rosenbloom KR, Raney BJ, Paten B, Nejad P, et al. The UCSC Genome Browser database: 2016 update. Nucleic Acids Res. 2015;44(D1):D717–25.

60. Ward LD, Kellis M. HaploReg: A resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 2012;40(D1):1–5.

61. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub M a., Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012;22(9):1790–7.

62. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 2014;42(D1):1001–6.

Supporting Information

S1 Fig. Retinal fundus image. Solid lines (red for arterioles and dark blue for venules) represent the vessels detected automatically and measured by VAMPIRE (Vasculature Assessment and Measurement Platform for Images of the REtina) software (version 3.0, Universities of Edinburgh and Dundee, UK). Dotted lines (light blue) represent the measurement zones on a fundus image; based on optic disc (light blue circle) location and radius.

S2 Fig. Meta-analysis on genome-wide association results from two independent discovery cohorts. Manhattan plots for six quantitative retinal traits. A. represents the results from Optic Disc Radius (*ODradius*), B. represents the results from retinal arteriolar tortuosity maximum

(*TortAmax*), C. represents the results from retinal venular tortuosity (*TortVmax*), D. represents the results from Central Retinal Arteriolar Equivalent (*CRAE*), E. represents the results from Central Retinal Venular Equivalent (*CRVE*), and F. represents the results from Arteriole-to-Venule ratio (*AVR*). The blue and red horizontal lines indicate the suggestive and genome-wide significance threshold ($P<5\times10^{-8}$), respectively.

S3 Fig. Quantile-quantile plots of GoDARTS-ORCADES meta-analysis for eight quantitative retinal blood vessel traits. Shaded areas represent 95% confidence intervals. Naturally log transformed -TortA: retinal arteriolar tortuosity, TortAmax: maximum retinal arteriolar tortuosity, TortV: retinal venular tortuosity, TortVmax: maximum retinal arteriolar tortuosity. ODradius: Optic Disc Radius, CRAE: Central Retinal Arteriolar Equivalent, CRVE: Central Retinal Venular Equivalent, AVR: Arteriole-to-Venule ratio.

S4 Fig. Regional association plots of index SNP reached p-value < $5x10^{-7}$ in the meta-analysis of the two discovery study cohorts (GoDARTS and ORCADES). a) *TortA* b) maximum *TortA* c-d) *TortV* e) *ODradius* f) maximum *TortV*.

S5 Fig. Box plots represent the expression level of genes associated with retinal blood vessel traits. a) *COL4A2* b) *COL4A1* c) *ACTN4* d) *CAPN12.* These plots were created using GENEVESTIGATOR which integrates manually curated gene expression data from microarray and RNAseq experiments. Blue lines at the bottom of the box indicates the gene expression level across 451 tissues in human. Y-axis depicts the top ten tissues and the sample size shown in secondary Y-axis. Colour scale (Low to High) at the top depicts the gene expression range in log2scale.

S6 Fig. eQTL annotation of top GWAS variants for quantitative retinal vessel traits. Y-axis represents tissue-specific gene expression data which is normalized by rank normalization method while x-axis shows the GWAS lead variant genotypes. a) TortA associated SNP, rs7991229 is correlated with *COL4A2* expression in heart left ventricle; b) TortA associated SNP, rs9515212 is correlated with *COL4A2* expression in heart left ventricle tissue; c) TortV associated SNP, rs1808382 is correlated with *CAPN12* expression in artery aorta; d) TortV associated SNP, rs1808382 is correlated with *ACTN4* expression in artery aorta.

S7 Fig. Conditional analysis of the genome-wide significant variant (rs56399312) at *COL4A2* locus. Locus zoom plots for the *COL4A2* locus associated region (GoDARTS) conditioned on the CAD associated SNPs (rs11617955, rs4773144, rs9515203) reported previously in the

748 GWAS study. a) Top SNP (TortA) Conditioned on rs11617955, b) Top SNP (TortA)
749 Conditioned on rs4773144 c) Top SNP (TortA) Conditioned on rs9515203.

750 **S8 Fig.** Multi-tissue eQTL comparison for TortV associated SNP, rs1808382 is correlated
751 with a) *ACTN4* expression and b) *CAPN12* expression.

752 **S9 Fig.** Locus zoom plot for the *ACTN4* locus associated with TortV also associated with
753 pulse rate in UK Biobank. The lead SNP, rs1808382 associated with TortV (Discovery and
754 replication stage) in that region is indicated by purple colour solid diamond.

755

756 **S1 Table.** Significant SNPs for each quantitative retinal vascular traits that reached $P<7x10^{-7}$ in
757 the meta-analysis of discovery cohorts.

758 **S2 Table.** SNPs in *COL4A2* conditioned on top SNP rs56399312 (discovery cohorts), associated
759 with *TortA*. TortA, retinal arteriolar tortuosity.

760 **S3 Table.** Summary of previously reported significant SNPs associated with Optic Disc area,
761 CRAE, and CRVE look-ups in GoDARTS-ORCADES meta-analysis study.

762 **S4 Table.** *Insilico* functional annotation of significant SNPs for quantitative retinal vascular
763traits.

764 **S5 Table.** Significant top hits that reached $P<1x10^{-7}$ for quantitative retinal vascular traits as
765 eQTLs (GTEx) in different tissues.

766 **S6 Table.** Top SNPs in *COL4A2* associated with TortA conditioned on reported coronary artery
767 disease SNPs.

768 **S7 Table.** Summary of significant SNPs ($P<8x10^{-07}$) associated with tortuosity traits from
769 discovery stage , replicated in myocardial infarction (MI) and coronary artery disease (CAD)
770 GWAS.

771 **S8 Table.** Summary of previously reported significant SNPs associated with coronary artery
772 disease (CARDioGramplus C4D) look-ups in GoDARTS-ORCADES meta-analysis study for
773 retinal arteriolar tortuosity trait.

774 **S9 Table.** Genome-wide significant SNPs ($P<8x10^{-07}$) for retinal tortuosity traits from discovery
775 stage, are associated with different traits (Type2Diabetes Knowledge Portal and ICBP).

776 **S10 Table.** Sensitivity analyses using GoDARTS (diabetes).

777 **S11 Table.** Lead SNPs associated with TortV are also associated with heart rate in UK Biobank.

Author Contributions

The study was designed by C.NA.P, A.SF.D, and E.T for GoDARTS cohort, J.F.W for ORCADES cohort, I.J.D for LBC1936 cohort, O.P for Croatia-Split, and Croatia-Korcula cohort. VAMPIRE software was designed and developed by E.T, T.M, D.R, E.B, S.K.V, and B.D. Retinal images were collected and analysis was performed by E.T, T.M, J.F.W, L.B, M.K, D.R, V.V, and H.C. Genotype data processing and statistical analysis was conducted by A.V, K.E.S, P.K.J, L.B, M.K, S.H, V.V, C.H and K.Z. Bioinformatics analysis was performed by A.V. The manuscript was drafted by A.V, C.NA.P A.SF.D, and revised by E.T, J.F.W, T.M, I.J.D, S.H, E.R.P and K.Z. All the authors reviewed the manuscript.

Figure Legends

Fig 1. Study Design. GoDARTS: Genetics of Diabetes Audit and Research in Tayside; ORCADES: Orkney Complex Disease Study; LBC1936: Lothian Birth Cohorts 1936; All Croatia: Croatia island of Korcula, Croatia Split; *ODradius*: Optic Disc Radius, *CRAE*: Central Retinal Arteriolar Equivalent, *CRVE*: Central Retinal Venular Equivalent, *AVR*: Arteriole-to-Venule ratio, Natural log transformed data – *TortA*: retinal arteriolar tortuosity, *TortAmax*: maximum retinal arteriolar tortuosity, *TortV*: retinal venular tortuosity, *TortVmax*: maximum retinal arteriolar tortuosity; PC: Principal Components; u is the genetic value for each subject under a random effects model, covariance amongst subjects assumed to be proportionate to the genomic relationship matrix.

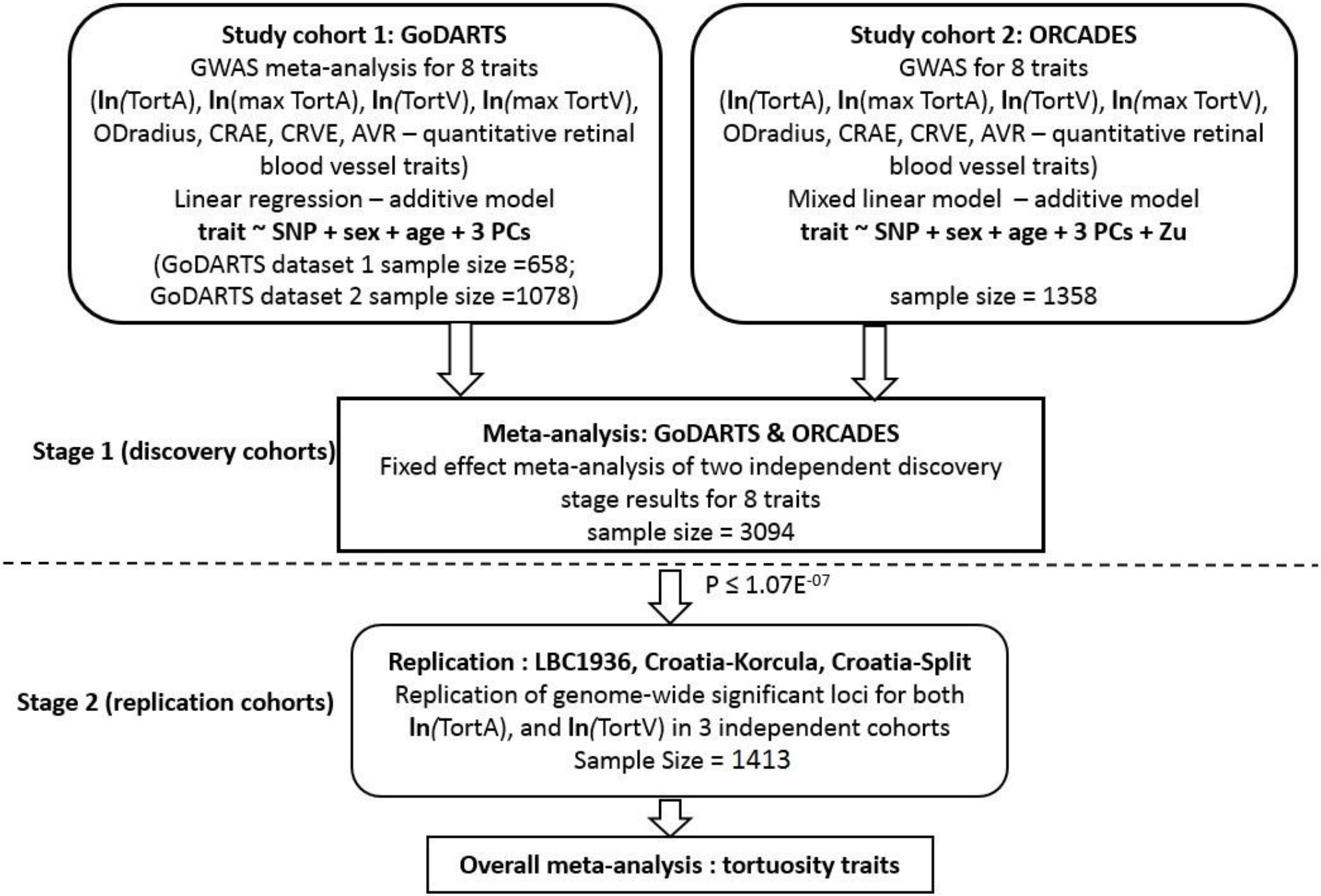
Fig 2. Manhattan plots for meta-analysis of genome-wide association results from two independent discovery cohorts. A, represents the results for the arteriolar tortuosity (*TortA*) and B. represents the results for the venular tortuosity trait (*TortV*). The blue and red horizontal lines indicate the suggestive and genome-wide significance threshold ($P<5\times10^{-8}$), respectively.

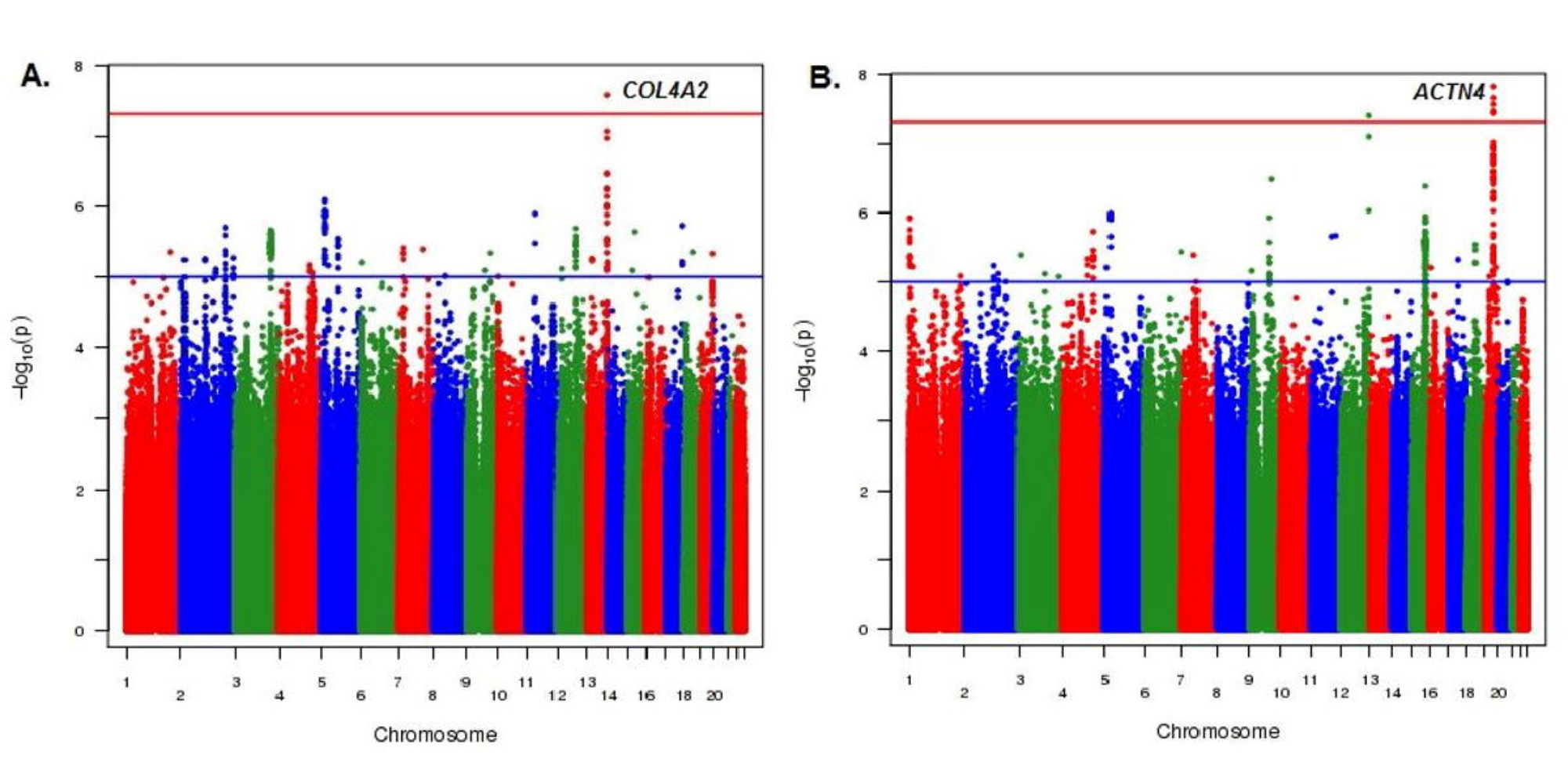
Fig 3. Regional association and recombination plots of variants that reached genome-wide significance in overall meta-analysis (discovery and replication stage). A-B, Top hits for *TortA* and C-D, Top hits for *TortV*. Each plot was created using LocusZoom for the lead SNP in genomic region 400 kb in either side of the significant signal. Blue spikes

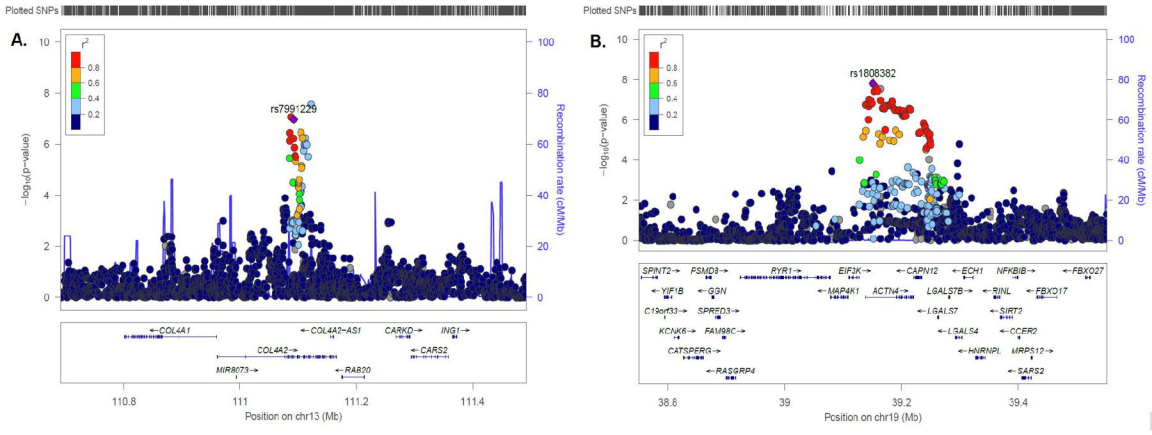
represents the estimated recombination rates. Colour scale (high to low r2) circles depicts the pairwise correlation (r2) between lead SNP and other SNPs in the loci. The lead SNP in that region is indicated by purple colour solid diamond and gene annotations in this region is shown in the bottom panels.

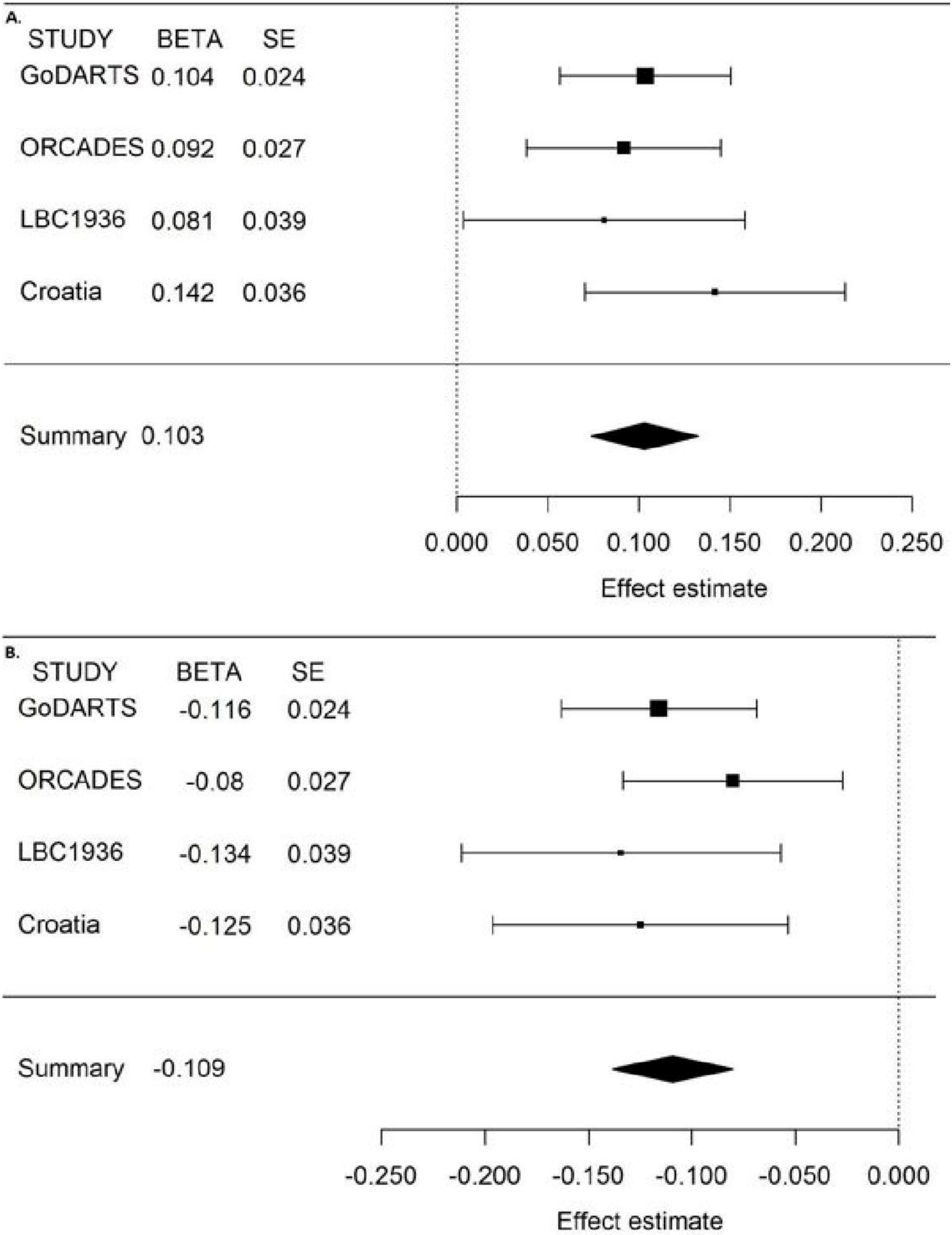
Fig 4. Forest Plots for the genome-wide significant hits (overall meta-analysis) associated with arteriolar (A. rs7991229) and venular tortuosity traits (B. rs1808382). The plots represent standardized beta and standard error from GoDARTS, ORCADES, LBC1936, Croatia KORCULA-SPLIT, and meta-analysis study. Standardized beta estimate: Change in natural log transformed retinal tortuosity traits for each copy of the effect allele. Due to the difference in the units of the beta and standard errors between the discovery and replication studies arising from different approaches to measurement, we standardized the effect estimates from each individual's study results.

820


Study cohort 1: GoDARTS
GWAS meta-analysis for 8 traits
(ln(TortA), ln(max TortA), ln(TortV), ln(max TortV), ODradius, CRAE, CRVE, AVR – quantitative retinal blood vessel traits)
Linear regression – additive model
trait ~ SNP + sex + age + 3 PCs
(GoDARTS dataset 1 sample size =658;
GoDARTS dataset 2 sample size =1078)
Study cohort 2: ORCADES
GWAS for 8 traits
(ln(TortA), ln(max TortA), ln(TortV), ln(max TortV), ODradius, CRAE, CRVE, AVR – quantitative retinal blood vessel traits)
Mixed linear model – additive model
trait ~ SNP + sex + age + 3 PCs + Zu
sample size = 1358
Stage 1 (discovery cohorts)
Meta-analysis: GoDARTS & ORCADES
Fixed effect meta-analysis of two independent discovery stage results for 8 traits
sample size = 3094
P ≤ 1.07E-07
Stage 2 (replication cohorts)
Replication : LBC1936, Croatia-Korcula, Croatia-Split
Replication of genome-wide significant loci for both ln(TortA), and ln(TortV) in 3 independent cohorts
Sample Size = 1413
Overall meta-analysis : tortuosity traits


A.
COL4A2
−log10(p)
Chromosome
B.
ACTN4
−log10(p)
Chromosome


Plotted SNPs
A.
r²
0.8
0.6
0.4
0.2
rs7991229
−log10(p-value)
Recombination rate (cM/Mb)
COL4A1
COL4A2
COL4A2−AS1
CARKD
ING1
CARS2
MIR8073
RAB20
110.8
111
111.2
111.4
Position on chr13 (Mb)
Plotted SNPs
B.
r²
0.8
0.6
0.4
0.2
rs1808382
−log10(p-value)
Recombination rate (cM/Mb)
SPINT2
FSMD8
RYR1
EIF3K
CAPN12
ECH1
NFKBIB
FBXO27
YIF1B
GGN
MAP4K1
ACTN4
LGALS7B
RINL
FBXO17
C19orf33
SPRED3
LGALS7
SIRT3
KCNK6
FAM98C
LGALS4
CCER2
CATSPERG
HNRNPL
MRPS12
RASGRP4
SARS2
38.6
39
39.2
39.4
Position on chr19 (Mb)


A.
STUDY BETA SE
GoDARTS 0.104 0.024
ORCADES 0.092 0.027
LBC1936 0.081 0.039
Croatia 0.142 0.036
Summary 0.103
0.000 0.050 0.100 0.150 0.200 0.250
Effect estimate
B.
STUDY BETA SE
GoDARTS -0.116 0.024
ORCADES -0.08 0.027
LBC1936 -0.134 0.039
Croatia -0.125 0.036
Summary -0.109
-0.250 -0.200 -0.150 -0.100 -0.050 0.000
Effect estimate